

TEACHER EVALUATION: PRINCIPALS' INSIGHTS AND SUGGESTIONS FOR IMPROVEMENT

Introduction

In the twenty-first century, school improvement is the focal point for educational leadership. Improvement must take the form of higher student achievement test scores, an engaging classroom that meets the needs of all learners, and a more meaningful learning environment for teachers and students. In the twenty-first century, the school building administrator is recognized as the catalyst for this necessary improvement. The school building administrator is no longer viewed as a manager, though s/he still has managerial roles. Rather the principal must be the instructional leader (Glickman, Gordon, & Ross-Gordon, 2003; Sergiovanni, 1994).

Within the area of teacher evaluation, the evaluation tools, along with the amount of time, format, and feedback, have changed dramatically from the past procedures used by the principal as manager (Danielson & McGreal, 2000; Schmoker, 2001). Two drop-in visits with a pre-made checklist are no longer considered acceptable practice (DeMoulin, 1988; Edmonds, 1981). The literature now considers the use of pre- and post-observation conferencing, narratives, rubrics, and portfolios as best practice procedures within teacher evaluation, with the result being school improvement (Danielson, 2002; Darling-Hammond, 1997; Kerrins & Cushing, 2000; Klein, 1990; Tucker & Strong, 2001). But do these practices really deliver? Is the time that is necessary to conduct these measures reasonable? Can one building principal do justice to these measures for all of their teaching staff? Do principals feel that the time and effort result in increased student learning? Are there still barriers to improving teaching and learning within the teacher evaluation process? Do principals believe that there is a better way to improve teaching and learning? These are the questions that motivated this study.

Literature Review

The demands for improving the quality of teaching and learning in public schools are as strong today as they were when the Soviet Union first launched Sputnik in 1957 or *A Nation at Risk* was published in 1983. Public school administrators and, in particular, school principals typically feel tremendous internal and external pressure to improve the quality of teaching in their schools in an effort to increase student achievement (Connors, 2000; Edmonds, 1981; Green, 2001; Schmoker, 2001). Internal pressure may result from the passion that many administrators bring to their positions or from the accountability demands built into collective bargaining agreements and/or state laws which can limit an administrator's authority within the evaluation process (Bolman & Deal, 2003; Marczyk & Marczyk, 2002; Murphy, 2002).

External pressure often emanates from a variety of sources including national and state-level politics, high profile business leaders,

and even local school district stakeholders who are not shy about voicing their concerns about the perceived lack of student achievement in public schools (Green, 2001; Senge, et al., 2000). In the last few years, this external pressure has been further exacerbated by the passage of the No Child Left Behind Act which imposes mandated sanctions for schools that do not meet increasing student achievement standards primarily related to standardized achievement tests (Illinois State Board of Education, 2004).

Research in educational leadership convincingly demonstrates that building level administrators are the central figures in the school improvement process (Murphy, 2002). Studies have consistently linked substantive student achievement gains to what happens at both the individual school and more specifically the classroom level (Danielson & McGreal, 2000; Fullan, 2001; Heller, 2004; Stigler, 1999). Since principals and assistant principals are directly responsible for faculty supervision, including teacher evaluation, it is not surprising that building administrators often find themselves held accountable for school improvement efforts and ultimately student achievement gains.

Within the Sputnik years, when schools were seen as needing to be improved, curriculum was the focus for fixing. A factory model was assumed within curriculum development resulting in programmed readers and independent learning modules in math and science. These were created by educational publishing companies whose “teams of experts” certified a consistent message and, therefore, a constant quality product—the child. These curricula were considered “fool-proof,” with the fool being considered the adults in the school. In dispensing this material, a good teacher was one who demonstrated certain traits: voice, appearance, accuracy, enthusiasm, and warmth, making certain that the children moved through the curriculum (Meux, 1974). The authors remember as children themselves moving from aqua to silver to gold in the reading series, with this color progression indicating that we were successful students who had learned the required curriculum. Therefore, the building principal, as manager, was to ensure that children progressed from one level to the next and that teachers were keeping accurate records of this progress and exhibiting the appropriate traits to foster it.

Moving from the 1960s into the '70s and '80s, the research focus shifted from programmed curriculum and teacher traits to an examination of necessary teacher skills and then to the Madeline Hunter model of teaching. Rather than prescribed curriculum, a prescribed delivery model was now seen as the road to school improvement. The building leader was to evaluate the teacher's performance against the prescribed teaching model. Principals were armed with checklists to verify that a participatory set, a mini-lesson, and guided practice were provided (Hunter, 1982). Cooperative learning, with its structures, was also seen as a way to improve student learning (Slavin, 1994). Again, the building principal was armed with a checklist to note simultaneous interdependence, student roles, and individual productivity. Throughout the late '60s, '70s and early '80s, the principal was busy checking off discrete items within a cookbook of school improvement. Professional development provided by both

the teacher unions and the administration focused upon the performance of these discrete measures. Teacher unions trusted these quantifiable checklists, since they rendered teaching and learning into black and white indices. Subsequently, teacher evaluation focused solely on these objective measures and it became the duty of the administrator to prove when the teacher did not demonstrate them (DeMoulin, 1988; Popham, 1992; Marczely & Marczely, 2002).

As the 1980s closed with the beginning of brain research, a more complex picture of teaching and learning began to emerge. From the late 1980s, through the '90s and now in the twenty-first century, ideas concerning teaching for understanding, authentic pedagogy, backwards design of curriculum, and the standards-based movement dominate the discussion of what is teaching and learning (Brophy, 1992; Haefele, 1993; Wiggins & McTighe, 1998). A mere checklist cannot capture the nuances of this new view of the teaching and learning process. This new conceptual model requires depth of understanding with multiple measures of academic and teaching performance. These multiple measures of teaching performance speak to the need for a more complex teacher evaluation model (Danielson, 2002; Rebore, 2000; Seyfarth, 2001). Additionally, the literature supports the belief of many stakeholders that improving the teacher evaluation processes will lead to instructional improvement and ultimately student achievement gains (Senge, et al., 2000; Tucker & Strong, 2001). Thus, a teaching evaluation model has emerged that is more detailed and takes more time and effort for both the teacher and the building administrator.

This shift in the research attracted the attention of state legislatures, many of which have passed legislation increasing teacher evaluation requirements and mandating teacher evaluation training for certified administrators. For example, in Illinois statewide legislated reform initiatives in the mid-eighties required school districts to develop specific teacher evaluation plans, including prescribed components, which were then submitted to the Illinois State Board of Education (ISBE) for approval. This legislation subsequently mandated that all administrators who evaluate teachers participate in a common state-approved multi-day workshop on teacher evaluation (ISBE, 1985, 1988, 2003). It was later amended to include annual, mandatory administrator attendance at state-approved workshops linked to school improvement—often focused on teacher evaluation (ISBE, 1995, 2001).

With these legislative changes, educators have watched teacher evaluation processes evolve from simple end of the year checklists or summative narratives to more sophisticated clinical processes and reflective teacher evaluation models (Danielson, 2002; Heller, 2004; Peterson, Wahlquist, Bone, Thompson, & Chatterton, 2001). Simultaneously, building-level administrators have experienced a dramatic increase in both the amount of time and skill required to implement these teacher evaluation processes while balancing the ever-increasing daily demands of administration (ISBE, 2002). Within this context of increased responsibilities and accountability several important concerns related to teacher evaluation,

school improvement, and increased student achievement remain unanswered. This research study seeks to explore these issues through the eyes of Illinois building-level school administrators who serve in the northern Cook County suburbs. While the research questions posed are asked from a broad perspective, the data are from a limited sample. Thus, generalization of the data to the entire state of Illinois or to the profession as a whole is subject to the interpretation of the researchers and the reader.

Problems and Purpose

As the emphasis on teacher evaluation has increased, so have both the number and variations of teacher evaluation models. In addition to the traditional summative evaluation that is characterized by checklists and/or narratives built around discrete performance indicators, the most prominent models are personal and instruction goal setting, clinical observation processes, portfolio assessment, and self-evaluation (Danielson, 2001; Kerrins & Cushing, 2000; Sawyer, 2001). Given the plethora of approaches, two significant questions are: Which teacher evaluation approaches are most widely employed in K-8 public education today? Do school districts employ the same or different teacher evaluation approaches with their tenured and non-tenured faculty?

In addition to identifying the types of teacher evaluation instruments employed, their effectiveness in improving teaching and learning is also an important consideration particularly since school-level administrators could potentially spend a great deal of their time completing the required teacher evaluation process (McGrath, 2000). As first line evaluators charged with improving the quality of teaching and student learning in their schools, these building-level administrators can provide valuable insights into the effectiveness of teacher evaluation as a tool. As such, their judgments about the impact of the various approaches to teacher evaluation are important in understanding more thoroughly the school improvement process (Wilkerson, Mannatt, Rogers, & Maughan, 2000). Therefore, this study seeks to understand the following: Do school building administrators perceive certain evaluation approaches to be more effective than others?

As the expectations for evaluating teachers have increased and the models employed expanded and become more sophisticated, administrators are expected to make teacher evaluation a priority (DeMoulin, 1988). Given the complexity of building-level administrators' responsibilities and the ever increasing demands on their time, this study asks: How much time do administrators spend completing all elements of their school district evaluation process with non-tenured and tenured faculty?

Additionally, teacher evaluation systems are continually evolving and increasing in importance within the teaching/achievement debate (Popham, 1992; Tucker & Strong, 2001). In practice, the present approaches to teacher evaluation may or may not be judged by school administrators to have positive impacts on teaching and learning. However, the process of teacher evaluation itself may yield other necessary benefits beyond the

improvement of teaching and learning emphasis. Simultaneously, because building administrators are intimately knowledgeable about their district-required teacher evaluation systems, they are in the unique position to identify specific policies, practices, and beliefs which inhibit optimal teacher evaluation effectiveness. Therefore this study seeks to investigate the following: What do building administrators perceive as benefits associated with teacher evaluation? What impediments do building-level administrators identify which may limit the effectiveness of the teacher evaluation system in their schools?

Finally, as principals and assistant principals go about their daily work as educational leaders, they have a unique understanding of most aspects of their individual school program and culture (Murphy, 2002). Because they work with faculty and students, provide leadership in curriculum and staff development, and interact with a multitude of internal and external stakeholders, administrators have the potential to understand what strategies might make a difference in improving teaching and student achievement (Senge, et al., 2000; Sergiovanni, 1994). These insights honed on years of teaching and administrative experience could prove very valuable in the improvement process (Glickman, et al., 2003). Therefore, this study seeks to understand administrative reflections pertaining to the following question: What do building administrators believe are the most valuable tools to improve teaching and student learning?

The Research Study

Context

Illinois is a state of contrasts. From the southern-most small town of Cairo to the northern metropolis of Chicago, Illinois is a state comprised of both cornfields and big cities. It is also “a state renowned for its legacy of local control of its public school districts” (Shanahan, 1988, p. 4). About 2,044,540 students are enrolled in 3,907 public schools in 887 districts within 110 counties grouped into six regions of the state (ISBE, 2004). According to the Illinois 2003–2004 State Report Card, these 2,044,540 children are 58.6% white, 20.7% black, 17% Hispanic, 3.6% Asian/Pacific Islander, and .2% Native American. Of these children, 37.9% come from low-income households, 6.3% come from homes whose native language is not English, and 14.8% attend schools that are on “School Improvement Status.”

This research study, conducted in May 2004, surveyed K–8 elementary and middle school building administrators in one section of one county in Illinois—suburban Northern Cook. Suburban Northern Cook County is bordered by DuPage County on the west, Lake County to the north, the City of Chicago and Interstate 80 to the south, and Lake Michigan to the east. In comparison with state totals, the racial make-up of the students within suburban Northern Cook County comprises fewer black children but more Asian/Pacific Islanders and approximately equal numbers of children who come from homes where English is not spoken.

Within Northern Cook County, 148,000 students are enrolled within 200 public schools. Of these 200 schools, 192 educate children in grades K–8. Administrators of these 192 schools, which compose 32 elementary school districts, were the target population for this study.

North Cook County Illinois was chosen for the study by the researchers because of the need for a high response rate from a limited population. Limited funding for the project deterred the researchers from surveying all of Cook County let alone all of Illinois. Therefore, a high response rate from the potential participants was desired. Both researchers had worked as school administrators in North Cook County and felt that their name recognition would motivate busy school administrators to respond resulting in a higher response rate for the study.

Method

Participants. Funding limitations dictated that a sample of 102 schools from the total 192 K–8 buildings be sent a survey addressed to the building administrator about the current system of teacher evaluation in the local school/school district and its perceived link to student performance. Within this sample, all thirty-two elementary school districts within North Cook County were represented by at least one school building. After ensuring that all districts were represented by one school building, the remaining schools were numbered and blindly selected through a rotation model. This was to make certain that larger school districts were represented in proportion to their relative size while ensuring that school buildings were chosen by chance rather than because of a possible special relationship with either researcher. Of the 102 surveys, 63 were completed and returned resulting in a 62% return rate.

Questionnaire. An 18-item self-administered questionnaire was developed and tested with a focus group of school administrators who were not to be included in the study. Once refined, the questionnaire and procedures were approved by a university Internal Review Board (IRB) and mailed to the aforementioned participant sample. Item development was based upon leadership, school improvement, teacher evaluation, and human resources literature and context-specific issues. Part One of the questionnaire asked building administrators to provide demographic data about their school district and school building along with their number of years as an administrator (see Appendix). Additionally, on the second page of the questionnaire, building administrators were asked to record the number of teachers they evaluated that year and the average amount of time they spend per year on non-tenured vs. tenured teacher evaluations.

Part Two of the questionnaire asked principals to indicate the tools they used for non-tenured and tenured teacher evaluation and to rate the effectiveness of these tools 1 through 5, with 1 as very high and 5 as no impact. Based on teacher evaluation literature (Danielson, 2001, 2002), the following tools were listed: end of year summative checklist authored by the principal, end of year summative narrative authored by

the principal, pre-observation conference, observation checklist, post-observation conference, portfolio authored by the teacher, and other. Within the category "other," space was provided for the respondent to describe the tool.

Part Three consisted of three open-ended items. Participants were first asked to describe the perceived benefits of their current evaluation system. Another item asked participants to describe the perceived impediments to this current teacher evaluation system. Finally, participants were asked to describe the activities they do, or would like to do, that they believe have, or would have, the greatest impact on teaching and learning.

Data collection. A hybrid of Dillman's (2000) Tailored Design Method and Queeney's (1995) Survey Design Method was used to maximize survey response. A cover letter, questionnaire, and addressed stamped envelope were sent to all participants. A post-card reminder was mailed approximately three weeks after the questionnaire. Additionally, on both the cover letter and the postcard, the phone numbers and email addresses of both researchers were provided, in case clarification was needed.

Data analysis. The survey data were entered in Microsoft Excel 2002 edition for analysis. Frequencies and percentages were used to describe closed-ended survey responses. Since participants ranged in years of experience, size of building, and number of non-tenured vs. tenured teachers to evaluate, baseline response expectations were not formulated. Instead, data were analyzed to identify any trends that might appear within these contextual categories (Maxwell, 1996). Through inductive analysis (McMillian & Schumacher, 2001) "categories and patterns emerged from the data gathered rather than being imposed apriori" (p. 462). Inter-data analysis was then used with these contextual trends to understand if any relations occurred between the context of the participant, the tool of evaluation, and the perceived effectiveness of the tool (Miles & Huberman, 1994). Additionally, quantitative data were then used as a reference point to verify and triangulate the open-ended feedback (Denzin, 1989). Through data reduction, data display, conclusion creation, and triangulation, verification and trend identification of the data were achieved (Berkowitz, 1997).

The data were initially analyzed and confirmed with the existing literature independently by each researcher in order to perform data reduction, display, and triangulation, and conclusion creation. Only after the entire inductive process was completed independently by each researcher did the two researchers then share their data-generated and literature-verified themes with each other. The researchers independently generated the same thematic concepts from the data for all of the results detailed in this paper. While this does not guarantee reliability and validity, it does provide "dependable results" (Guba & Lincoln, 1981, p. 146) which create a "rich and thick description of the context being examined and which can be retested by either quantitative or qualitative means to further establish

reliability and validity” (Merriam, 1988, p. 19). To that end, we have appended a copy of the survey instrument for further replication.

Furthermore, because themes were generated inductively after data were collected, some data that were gathered did not prove relevant after the inter-data analysis. In particular, the inductive process with subsequent inter-data analysis rendered information concerning the demographics of the school building, the number of years as administrator, and the number of teachers evaluated as non-explanatory themes for this data set. Rather the theme of non-tenured vs. tenured consistently appeared as an organizing element for the data.

Results

Teacher Evaluation Approaches and Their Effectiveness

School administrators report that post-observation conferences, pre-observation conferences, and summative narratives were the most widely used evaluation tools for both tenured and non-tenured staff faculty members. Percentages equal more than 100% since most administrators use a combination of tools for both non-tenured and tenured staff. The largest variation between the two faculty groups occurred within the areas of pre-observation conference and other. The data indicate that school administrators often will forgo pre-observation meetings with tenured faculty as well as provide alternative evaluation opportunities for these staff members. These other opportunities included the following: research action plans, grade level collaboration projects, and teacher leadership responsibilities including mentoring, coaching, and committee stewardship.

Table 1

Teacher Evaluation Approaches Employed

Approach	% of teachers who used approach	
	Non-tenured	Tenured
Summative checklist	42	43
Summative narrative	85	83
Pre-observation conference	91	75
Observation checklist	43	32
Post-observation conference	98	85
Portfolio	25	26
Other	25	38

School administrators were asked to rate each of the tools mentioned in Table 1 on a closed matrix scale from 1–5, with 1 as very highly effective and 5 as not effective at all. Averages were then compiled from these responses (see Table 2). Rating averages clustered around the 2.5 mark indicating a limited degree of effectiveness with the teacher evalua-

tion tools.

Table 2

Average Effectiveness Ratings of Teacher Evaluation Approaches

Approach	Average effectiveness rating
Summative checklist	2.55
Summative narrative	2.68
Pre-observation conference	2.57
Observation checklist	2.58
Post-observation conference	2.2
Portfolio	2.61
Other	2.5

Time Required

Table 3 details the responses by the school administrators about the average amount of time devoted to the annual evaluation of both individual non-tenured and tenured teachers. For non-tenured teachers, over 50% of school administrators spend more than ten hours per teacher each year on the full evaluation process including preparation, observation, conferencing, and documentation. For tenured teachers, 64% of administrators spend five hours or more per teacher on the evaluation process.

Table 3

Time Spent on the Evaluation Process

Time spent	% of building administrators who spend this amount of time on . . .	
	Non-tenured faculty	Tenured faculty
Up to 2 hours	0	7
2–4 hours	6	30
5–7 hours	19	30
8–10 hours	19	17
More than 10 hours	55	17

Teacher Evaluation Benefits

As study data demonstrate, school administrators devote a considerable portion of their time to implement the district required teacher evaluation process even though they generally perceive that this teacher evaluation process has a limited direct impact on improving an individual's teaching and subsequent student learning. However, the question

remains whether or not building level administrators perceive any other benefits associated with the teacher evaluation process. Administrators identified four areas of benefit associated with teacher evaluation, some listing more than one benefit.

Goal setting. The most frequently reported benefit (31%) of teacher evaluation was its usefulness in aligning individual teacher's personal goal setting with annual district goals. The inclusion of goal setting components in the teacher evaluation process helped administrators to work with faculty members to develop well-aligned individual annual goals that often included action plans. Another benefit of the goal setting process was that it contributed to richer discussions between the evaluator and teacher, particularly with tenured faculty, and that as a result of this discussion personal staff development plans were often created. Others noted the value of goal setting in assisting non-tenured teachers to be more reflective about their teaching.

Enhanced supervision. A significant percentage of respondents (26%) indicated that their current teacher evaluation processes, when administered well, had an impact on the efficacy of the supervision process. Most frequently identified benefits of this enhanced supervision included:

- Delineating specific focuses for classroom observations which in some instances targeted teaching deficiencies while providing a structure for the overall process;
- Providing a vehicle for the development of in-depth observation narratives which were used to set higher performance expectations;
- Identifying teaching strengths and weaknesses, particularly with non-tenured faculty; and,
- Providing an option to extend teacher evaluation beyond traditional classroom observation and summative narratives into professional development projects.

Enhanced communication. Another perceived positive of current teacher evaluation was enhanced communication. Administrators (23%) reported that their required plans created an effective structure for teacher/evaluator discussion. Benefits included:

- Establishing specified times for observation and pre- and post- conferencing that provided for focused communication;
- Creating a format for the sharing of expectations, school values, and curricular issues between the evaluator and faculty member;
- Developing a higher level of professional dialogue particularly with tenured teachers; and,

- Helping to create a common school-wide language related to teaching and learning.

Comprehensive process. The comprehensiveness of their district evaluation plans was mentioned as a benefit by 20% of the participants. In these districts, administrators reported that written, fully-developed standards and well-defined processes, often including evaluation rubrics and structured evaluation timelines, added substance to the overall model and led to more complete supervision and evaluation. In addition, some respondents noted that an important benefit of their comprehensive processes was that they were useful in holding faculty accountable for basic overall performance. Respondents listed this area of comprehensiveness separately from other tools that might be used such as the portfolio or action research model.

Impediments to Highly Effective Teacher Evaluation

A large percentage of participants (87%) identified one or more impediments that they perceived impacted their ability to implement fully the teacher evaluation process in their schools. Three primary impediments were reported: time, unions and school culture, and evaluation process constraints. These impediments were perceived by the building-level administrators to limit the effectiveness of their teacher evaluation processes. Some respondents indicated more than one impediment.

Time. Limitations associated with time were identified by 47% of the participants. This response rate indicates that building-level administrators believe that they do not have adequate time available to devote to teacher evaluation particularly since their teacher evaluation processes are time intensive and difficult to implement given the other demands of school administration. Many reported that they were required to evaluate too many teachers each year and that the paperwork demands associated with their district evaluation process were very extensive. Several participants felt that they did not have adequate time to supervise their non-tenured faculty who they believed needed more intensive supervision. In instances where school districts required the annual evaluation of all teachers, administrators reported that they found it necessary to shortcut parts of the process just to complete even the most basic elements. Finally, a lack of time to train administrators impeded the process and often resulted in inconsistent teacher evaluations within individual school districts.

Unions and school culture. Concerns related to the impact of unions and the pre-dominant culture of schools, in general, were reported by 36% of study participants as significant impediments to effective teacher evaluation. In a number of school districts collective bargaining agreements were found to limit the scope and substance of teacher evaluation. A consequence of union activity reported by some administrators was the willingness of faculty members to challenge some aspect of either the evaluation

procedures or content of their personal evaluations by taking their concerns directly to the local teacher unions/associations which often then intervened in the process. As a result, administrators noted a disincentive for thorough, honest evaluations and an erosion of confidence in the evaluation process.

This union role in collective bargaining, including in teacher evaluation, may be more symptom than cause. Over the past few decades, both the role and impact of unions have grown significantly in response to the increased sophistication of public education and the ever-expanding political nature of school governance, particularly related to faculty employment and retention. However, as discussed in the review of the literature concerning unions and teacher evaluation, unions were relatively comfortable with objective check-lists of teacher performance. Further research is necessary to understand whether the erosion in confidence in the evaluation process is due to the more subjective nature of increasingly more complex teaching evaluation tools being used or a function of the political nature of school governance.

In part because of these factors, the culture of public schools itself was viewed as an impediment. Building-level administrators freely discussed the prominent role that the culture of public education plays particularly within the teacher evaluation process. Administrators noted that teachers typically expect to receive excellent evaluations and resist evaluation methods that deviate from the status quo. Faculty willingness to experiment with something new in the evaluation process was often fraught with teacher resistance or passivity. Some administrators commented that they do not perceive school cultures embracing teacher evaluation as a tool for improvement but rather as a task which both the teacher and administrator are required to endure. This is contrary to the spirit of the new teacher evaluation tools that are considered to be best practice.

Evaluation process constraints. Finally, 14% of the administrators noted that they are required to implement an evaluation process that in and of itself lacks effectiveness. Predominant comments received included:

- The evaluation system is out of date and has not changed in decades;
- The system is not comprehensive enough to have any real impact;
- The criteria for ratings were inadequately defined and inconsistently interpreted; and,
- Although a district-wide process is in place, it does not yield any meaningful feedback for teachers.

Some noted that their processes were too summative, failing to provide any substantial emphasis on instructional improvement, while others said that evaluation procedures were either too vague or too generic to be of any substantive use. Some administrators saw this lack of perceived

effectiveness of their present teacher evaluation processes as the cause for teacher evaluation having no real impact on individual teacher behavior. And yet, when the evaluation tools were designed using best practice and the process was more formative than summative, teacher unions became uncomfortable and resisted this process too.

Improving Teaching and Learning

If the data suggest that building-level administrators perceive their present teacher evaluation processes and practices having minimal direct impact on improving teaching and learning, what do they believe would directly impact teaching and learning more? Four specific recommendations emerged from the data: increase communication, expand staff development, increase coaching and mentoring opportunities, and model classroom behaviors for teachers. Combined percentages exceed 100% as respondents often indicated more than one benefit.

Increased communication. If building-level administrators and teachers could be freed from the restraints created by time, as well as from union and school culture constraints, 69% believe that increased opportunities for communication would lead to significant improvements in teaching and learning. These administrators said that by increasing the opportunities for teachers to dialogue with administrators around the topics of teaching and learning, direct improvement could result. Although the role of building administrator as instructional leader is often articulated as a priority, the literature indicates that an evaluator's time is consumed by other responsibilities not related to improving teaching and learning (Edmonds, 1981; McGrath, 2000). By significantly increasing specific communication opportunities, building-level administrators believed that they could make a difference in teaching and learning. Respondents most frequently mentioned the following suggestions:

- Provide time for more one-on-one conversations between teachers and administrators, as well as between teachers, about teaching, learning, and the curriculum;
- Build into the evaluation system an emphasis on faculty self-reflection;
- Minimize summative and maximize formative evaluation processes;
- Structure opportunities for teachers to discuss student work and assessment;
- Encourage administrators to learn as much as they can about the day-to-day classroom activities and the instructional content through informal visitations;

- Build in frequent conversations with faculty-as-a-whole about school supervision and goal setting particularly as they relate to curriculum; and
- Expand the number of required classroom observations and increase the specificity of feedback.

Staff development. Twenty-three percent of the administrators believed that staff development would improve teaching and learning. While the responses could be interpreted as another form of communication, on further triangulation of the data, specific themes emerged within this category. Respondents most frequently mentioned:

- Tie staff development directly to student data;
- Tie staff development directly to formal growth plans;
- Make staff development consistent and focused on best practice; and
- Provide staff development time for grade-level collaboration.

Coaching and mentoring. Seventeen percent of administrators reported that formal coaching and mentoring programs would also directly improve teaching and learning. Building leaders noted that these coaching and mentoring opportunities should provide opportunities to:

- Model lessons;
- Co-teach;
- Meet as grade-level teams to create exemplary lesson plans; and
- Meet as colleagues for “book group” learning.

Modeling. Finally, 14% of administrators believed that when they have the time to actually demonstrate teaching to faculty, this modeling directly improves teaching and learning. Specifically, such modeling would focus on opportunities to:

- Model specific lessons focused on standards-based learning;
- Demonstrate specific instructional techniques; and
- Co-teach with faculty.

Limitations

This survey was distributed to school building-level administrators within Suburban North Cook County in the state of Illinois. While respondent rate was high, the generalizations made can only be validly applied to the specific context. Transfer of findings may be possible through replica-

tion of this survey study to the state or the nation as a whole. Additionally, survey data can be limiting in that respondents could display inconsistencies between their quantitative and qualitative responses. Through continual respondent feedback and data triangulation the researchers attempted to clarify these possible respondent discrepancies. (Denzin, 1989; Miles & Huberman, 1994.) However, caution must be exercised when drawing conclusions from such data.

Discussion

This study sought to understand building administrators' perceptions concerning the effectiveness of the teacher evaluation tools that school districts employ. School leaders, policy makers, board members, teacher union leaders, and classroom teachers may find the results useful in their quest for building evaluation systems that truly improve teaching and learning.

The data illustrate, first, that school administrators believe that evaluation is beneficial within the areas of goal setting and enhanced supervision and communication; and, second, that teacher evaluation systems have additional potential to benefit the teaching-learning process. Furthermore, when the system is comprehensive, the evaluation process can provide a focus and format for professional dialogue and the creation of a common school-wide language related to teaching and learning. These results are consistent with the current literature on effective teacher evaluation (Danielson, 2002; Schmoker, 2001).

In addition, the data strongly suggest that time, unions and school culture, as well as actual system constraints often negate many of the benefits of the teacher evaluation process. Respondents overwhelming (more than 50%) spend an enormous amount of time (more than 10 hours per teacher each year) on the evaluation process with non-tenured faculty. With such a large expenditure of time, it is understandable that completing the process could take precedence over meaningful conversations that could directly affect teaching and learning. This is contrary to what best practice in teacher evaluation should be. When understanding the potential benefits of teacher evaluation, a reexamination of what type, how much, and for whom, possibly along non-tenured/tenured lines, is necessary if we want best practices in teacher evaluation to be conducted in a meaningful, mindful manner.

Additionally, the complexity of public education and the growing political nature of school governance, particularly related to teacher unions and school culture, need to be examined further. Respondent data suggest that these factors contribute to a culture within schools that supports the status quo and squashes risk-taking and innovation. Additionally, system constraints play a part in impeding the promise of teacher evaluation. Further study needs to be conducted to understand the relationships among time, unions, politics, school culture, and system restraints to see if it is possible to create evaluation systems that are embraced by unions, benefit teaching and learning, and are not so cumbersome that school administra-

tors drown from the workload.

Finally, school administrators do believe they can make a difference in teaching and student learning. While noting the impediments, especially time, school administrators believe that, through increased communication opportunities, data-driven targeted staff development, peer coaching and mentoring, as well as principal demonstration teaching, they can improve instruction in the classroom. Many of these suggestions are supported in the current literature on school improvement (Sawyer, 2001; Senge, et al., 2000; Stigler, 1999). The concept of principal demonstration teaching suggests an even more active role for the school administrator as an instructional leader. However, if the issue of time is not addressed, the school administrator will not be able to take advantage of this possible school improvement strategy. The question is how to provide the means within teacher evaluation, as well as the time, for these types of meaningful opportunities to occur.

Conclusion

Teacher evaluation, when conducted appropriately, has the potential to improve teaching and learning. However, our research data suggest that, in our quest for a more comprehensive teacher evaluation system that depicts the true nature of teaching and learning, we may have created a monster. Our data indicate that principals believe that the current teacher evaluation systems are inordinately time intensive and preclude many other opportunities for school building leaders to work with faculty to improve classroom instruction. Additionally, our research data explain that principals perceive unions as not trusting the more complex, subjective teacher evaluation methods that are currently considered best practice. Therefore, further study must be done to understand the real potential of alternatives, especially within the areas of peer coaching and mentoring, that provide avenues for shared leadership, including the concept of principal as fellow and model teacher. Additionally, further research needs to be done to ascertain the strategies necessary to promote shared responsibility and trust between management and unions if the evaluation process is going to be a vehicle for school improvement and positive change.

References

- Berkowitz, S. (1997). Analyzing qualitative data. In J. Frechtling & L. Sharp (Eds.), *User-friendly handbook for mixed methods evaluations* (pp. 4–22). Arlington, VA: National Science Foundation.
- Bolman, L. G., & Deal, T. T. (2003). *Leading with soul*. San Francisco: Jossey-Bass.
- Brophy, J. (1992). Probing the subtleties of subject-matter teaching. *Educational Leadership*, 49, 4–8.
- Connors, N. A. (2000). *If you don't feed the teachers: They eat the students!* Nashville, TN: Incentive.

- Danielson, C. (2001, February). New trends in teacher evaluation. *Educational Leadership*, 58(5), 12–15.
- Danielson, C. (2002). *Enhancing student achievement: A framework for school improvement*. Alexandria, VA: ASCD.
- Danielson, C., & McGreal, T. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: ASCD.
- Darling-Hammond, L. (1997). *The right to learn*. San Francisco: Jossey-Bass.
- DeMoulin, D. F. (1988). Staff development and teacher effectiveness: Administrative concerns. *Focus*, 81, 36–38.
- Denzin, N. K. (1989). *The research act: A theoretical introduction to sociological methods* (4th ed.). Burr Ridge, IL: McGraw-Hill.
- Dillman, D. A. (2000). *Mail and internet surveys: The tailored design method* (2nd ed.). New York: Wiley.
- Edmonds, R. (1981). *Characteristics of schools that are instructionally effective for all pupils*. Alexandria, VA: ASCD.
- Fullan, M. (2001). *Leading in a culture of change*. San Francisco: Jossey-Bass.
- Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. (2003). *Supervision and instruction* (6th ed.). Needham Heights, MA: Allyn & Bacon.
- Green, R. L. (2001). *Practicing the art of leadership: A problem-based approach to implementing the ISLLC standards*. Upper Saddle River, NJ: Merrill/Prentice-Hall.
- Guba, E. G., & Lincoln, Y. S. (1981). *Effective evaluation*. San Francisco: Jossey-Bass.
- Haefele, D. L. (1993). Evaluating teachers: A call for change. *Journal of Personnel Evaluation in Education*, 7, 21–31.
- Heller, D. A. (2004). *Teachers wanted: Attracting and retaining good teachers*. Alexandria, VA: ASCD.
- Hunter, M. C. (1982). *Mastery teaching*. Thousand Oaks, CA: Corwin Press.
- Illinois State Board of Education (ISBE). (1985). *Illinois project for school reform*. Springfield, IL: Author. Available from <http://www.isbe.net>
- ISBE. (1988). *Overview and response to questions about teacher evaluation*. Springfield, IL: Author. Available from <http://www.isbe.net>
- ISBE. (1995). *Comprehensive school reform for Illinois' school children*. Springfield, IL: Author. Available from <http://www.isbe.net>
- ISBE. (2001). *Guidelines for implementing NCLB*. Springfield, IL: Author. Available from <http://www.isbe.net>
- ISBE. (2002). *Illinois Professional School Leader Standards*. Springfield, IL: Author. Available from <http://www.isbe.net>
- ISBE. (2003). *Illinois School Code 105 ILCS 5/24A-5*. Springfield, IL: Author. Available from <http://www.isbe.net>
- ISBE. (2004). *Illinois School Report Card 2004*. Springfield, IL: Author. Available from <http://www.isbe.net>

- Kerrins, J. A., & Cushing, K. S. (2000). Taking a second look: Expert and novice differences when observing the same classroom teaching segment a second time. *Journal of Personnel Evaluation in Education*, 12(10), 5–24.
- Klein, T. J. (1990). Performance reviews that rate an A. *Personnel*, 67(5), 38–41.
- Marczely, B., & Marczely, D. W. (2002). *Human resource and contract management in the public school: A legal perspective*. Lantham, MA: Scarecrow Press.
- Maxwell, J. A. (1996). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage.
- McGrath, M. J. (2000). The human dynamics of personnel evaluation. *School Administrator*, 57(9), 34–38.
- McMillan, J. H., & Schumacher, S. (2001). *Research in education* (3rd ed.). New York: Addison Wesley Longman.
- Meux, M. (1974). Teaching the art of evaluating. *Journal of Aesthetic Education*, 8(1), 85.
- Merriam, S. B. (1988). *Case study research in education: A qualitative approach*. San Francisco: Jossey-Bass.
- Miles, M. M., & Huberman, A. M. (1994). *Qualitative data analysis*. Thousand Oaks, CA: Sage.
- Murphy, J. (2002). Reculturing the profession of educational leadership: New blueprints. In J. Murphy (Ed.), *The educational leadership challenge: Redefining leadership for the 21st century* (pp. 65–82). Chicago: National Society of Education.
- National Commission on Excellence in Education. (1983). *A nation at risk*. Washington, DC: Author.
- Peterson, K. D., Wahlquist, C., Bone, K., Thompson, J., & Chatterton, K. (2001). Using more data sources to evaluate teachers. *Educational Leadership*, 58(5), 40–43.
- Popham, W. J. (1992). *Educational evaluation* (3rd ed.). Rockleigh, NJ: Allyn & Bacon.
- Queeney, D. S. (1995). *Assessing needs in continuing education: An essential tool for quality improvement*. San Francisco: Jossey-Bass.
- Rebore, R. W. (2000). *Human resources administration: A management approach* (6th ed.). Boston: Allyn & Bacon.
- Sawyer, L. (2001). Revamping a teacher evaluation system. *Educational Leadership*, 58(5), 44–47.
- Schmoker, M. (2001). *The results fieldbook: Practical strategies from dramatically improved schools*. Alexandria, VA: ASCD.
- Senge, P., Cabron-McCabe, N., Lucas, T., Smith, B., Dutton, J., & Kellner, A. (2000). *Schools that learn: A fifth discipline fieldbook for educators, parents and everyone who cares about education*. New York: Warner Books.
- Sergiovanni, T. J. (1994). *Building community in schools*. San Francisco: Jossey-Bass.
- Seyfarth, J. T. (2001). *Personnel management for effective schools* (3rd ed.). Needham Heights, MA: Allyn & Bacon.

- Shanahan, T. (1988). The voyage into school reform. *Illinois Schools Journal*, 68(1), 3–8.
- Slavin, R. E. (1994). *Cooperative learning theory: Research and practice* (2nd ed.). Needham Heights, MA: Allyn & Bacon.
- Stigler, J. W. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York: Free Press.
- Tucker, P. D., & Strong, J. H. (2001). Measure for measure: Using student test results in teacher evaluation. *American School Board Journal*, 188(9), 34–7.
- Wiggins, G., & McTighe, J. (1998). *Understanding by design*. Alexandria, VA: ASCD.
- Wilkerson, D. J., Manatt, R. P., Rogers, M. A., & Maughan, R. (2000). Validation of student, principal, and self-ratings in 360-degree feedback for teacher evaluation. *Journal of Personnel Evaluation in Education*, 12(2), 179–192.

Thomas A. Kersten is an Assistant Professor in the Department of Educational Leadership at Roosevelt University, Chicago, Illinois.

Marla S. Israel is an Assistant Professor in the Department of Educational Leadership at Loyola University, Chicago, Illinois.

Appendix

Teacher Evaluation Survey © Kersten & Israel 2005

Demographic Data:

Your Position: ___ Principal ___ Asst. Principal ___ Dept. Chair

Experience: Total years in school administration _____

Your Building: ___ Elementary
 ___ Middle School/Junior High
 ___ High School

District: ___ Elementary ___ High School ___ Unit

District Enrollment: ___ Less than 1,000
 ___ 1,000–3,999
 ___ 4,000 or over

Teacher Evaluation Tools & Effectiveness:

Please check all that are used. Next to each tool used, please note its effect on improving teaching and learning on a 1–5 Effectiveness Rating (ER) scale with 1 as very high and 5 as no impact.

Tool	Non-Tenured	Effectiveness Rating	Tenured	Effectiveness Rating
End of the Year Summative Checklist				
End of the Year Summative Narrative				
Pre-Observation Conference				
Observation Checklist				
Post-Observation Conference				
Portfolio				
Other *				

* Please describe _____

Additional Questions:

How many certified faculty members did you evaluate during the 2003–04 school year? _____

On the average, how much time per year do you spend completing all aspects (preparation, observation process, meeting, write up, documentation, etc.) for evaluating a typical non-tenured teacher?

_____ Up to 2 hours	_____ 8–10 hours
_____ 2–4 hours	_____ Over 10 hours
_____ 5–7 hours	

On the average, how much time per year do you spend completing all aspects (preparation, observation process, meeting, write up, documentation, etc.) for evaluating a typical tenured teacher?

_____ Up to 2 hours	_____ 8–10 hours
_____ 2–4 hours	_____ Over 10 hours
_____ 5–7 hours	

What do you see as the primary benefit of your district's present teacher evaluation system?

What do you see as the most significant impediment to teacher evaluation?

What do you do, or would you do, to provide the greatest impact on teaching and learning?

Thank you for your assistance.